



Софийски университет „Св. Кл. Охридски“

Факултет по математика и информатика

*Бакалавърска програма
„Софтуерно инженерство“*



Предмет: XML технологии за семантичен Уеб

Зимен семестър, 2017/2018 год.

Тема №08: „Каталог на университети - 2“

Курсов проект

Автори:

Ева Паунова, фак. номер 61940

Лиляна Видева, фак. номер 61887

януари, 2018

София

Съдържание

1. Въведение.....	3
1.1.Задание на проекта.....	3
1.2.Актуалност на избраната тема.....	3
1.3.Какъв проблем решава проекта.....	3
1.4.Какво използва проекта, за да реши зададения проект.....	3
1.5.Структура на останалата част от документа.....	4
2. Анализ на решението.....	4
2.1.Работен процес.....	4
2.2.Структура на съдържанието.....	5
2.3.Тип и представяне на съдържанието.....	7
3. Дизайн.....	7
4. Тестване.....	8
5. Заключение и възможно бъдещо развитие.....	10
6. Разпределение на работата.....	10
7. Използвани литературни източници и уеб сайтове.....	10

1 Въведение

1.1. Задание на проекта

Да се създаде каталог на университетите в България по региони, базиран на XML документи с текстово и графично съдържание, описващо възможните характеристики на всеки един университет в страната. Каталогът да представя графичното съдържание чрез XML Entities. Врзките в каталога – напр. Между региони и университети, между университети и техните филиали и др. – да се описват чрез (съставни) ключове и референции към тях. Да се състави описанието на 7 – 8 университета и да се валидира чрез подходящо изграден за целта XML Schema документ. Създаденото XML съдържание да се представи в HTML браузер чрез CSS и XSLT, с възможност за сортиране на представените обекти по няколко от техните признаци.

1.2. Актуалност на избраната тема за проект

При избор на университет е важно да имаш ясна и подробна информация относно наличните възможности, местоположение, какви факултети има, специалности, които съответният университет предлага. Вместо да търсиш например информацията за факултетите в сайта на всеки един университет по отделно, би било хубаво тази информация да бъде достъпна добре структурирана на едно място, и съответно да търсиш подробности само за отговарящите на твоите изисквания университети. Представеният проект фокусира точно върху този проблем.

1.3. Какъв проблем решава проекта?

Създаденият проект „*Каталог на университети*“, съдържа списък от най – големите университети в България. За всеки университет е представена главната информация – име, местоположение, регион, в който се намира, телефонен номер, ако е наличен – втори телефонен номер, година на създаване, сайт на университета, брой студенти, както и списък на факултетите. Каталогът предлага разглеждане на университетите, групирани по региони, както и разделени по филиали. Проектът съдържа информацията за 10 от най – големите университети в България, намиращи се във всичките 6 региона на страната, както и за техните филиали (ако има такива). Всички университети са сортирани в низходящ ред спрямо броя на студентите в тях.

1.4. Какво използва проекта, за да реши зададения проблем?

Проектът използва XML Entities, за да представи нужната информация за всеки един университет. Врзките между университети и региони, както и

между университети и техните филиали, се осъществяват чрез (съставни) ключове и референции към тях. Съдържанието на XML документа е валидирано чрез изградената XML Schema. За графичното представяне съдържанието в уеб браузър, се използва XSLT документ, който преобразува XML данните. Използвано е CSS стилизиране на съдържанието в уеб браузър.

1.5. Структура на останалата част от документа

В останалата част от документацията на настоящия проект са включени следните точки:

2. Анализ на решението

- Описание на входните документи, както и тяхното съдържание

3. Дизайн

- Какви са връзките, чрез които е изграден каталога, както и как е валидиран

4. Тестване

- Тестване на проекта чрез различни браузъри, както и резултатът от тестването

5. Заключение и възможно бъдещо развитие

6. Разпределение а работата

7. Използвани литературни източници и уеб сайтове

2 Анализ на решението

2.1 Работен процес

- Входното съдържание на каталога се съдържа в документа universityCat.xml. Информацията, както текстовото, така и графичното съдържание, е получено от уеб сайтовете на съответните университети - www.uni-sofia.bg, www.tu-sofia.bg, www.aubg.edu.
- Входното XML съдържание се валидира чрез XML Schema. За визуализацията в уеб браузър се използва XSLT трансформация, както и CSS стилизиране.
- Създаденият „Каталог на университети“ е лесно достъпен за потребителите в уеб браузър - Mozilla Firefox, Microsoft Edge и други.

```

<unisInBG>
  <universityList>
    <university>
      <uniInformation>
        <name>Медицински университет - Плевен</name>
        <location>Плевен</location>
        <region>Северен централен регион</region>
        <id>10</id>
        <regionId>201</regionId>
        <telephone>064/884-101</telephone>
        <telephoneExtra>064/800-728</telephoneExtra>
        <year>1974</year>
        <students>1458</students>
        <site href="http://www.mu-pleven.bg/">www.mu-pleven.bg</site>
      </uniInformation>
      <facultyList>
        <faculty>Факултет "Медицина"</faculty>
        <faculty>Факултет "Фармация"</faculty>
        <faculty>Факултет "Обществено здраве"</faculty>
        <faculty>Факултет "Здравни грижи"</faculty>
      </facultyList>
    </university>
  </universityList>
  <branchList>
    <branch>
      <seatName>Технически университет - София</seatName>
      <uni nameRef="Технически университет - София" idRef="2"/>
    </branch>
  </branchList>
  <regionList>
    <regionInfo>
      <regionName>Северозападен регион</regionName>
      <reg regRef="Северозападен регион" regIdRef="201"/>
    </regionInfo>
  </regionList>
</unisInBG>

```

Фигура 1: Извадка от съдържанието на XML документа

2.2 Структура на съдържанието

Каталогът има следната обща структура:

1. unisInBG – Коренов елемент, който съдържа следните елементи:
 - 1.1. universityList – съдържа списък на университетите; може да се съдържа най – много веднъж в unisInBG. Съдържа следните елементи:
 - 1.1.1. university – университет; може да се съдържа неограничен брой пъти в universityList. Съдържа следните елементи:
 - 1.1.1.1. uniInformation – информация за конкретния университет. Съдържа следните елементи:
 - 1.1.1.1.1. name – име на университета, задължителен елемент от тип String
 - 1.1.1.1.2. branchName – име на филиала, ако университетът е филиал на някой друг, незадължителен елемент от тип String

- 1.1.1.1.3. location – име на населеното място, в което се намира университета, задължителен елемент от тип String
- 1.1.1.1.4. region – име на региона, в който се намира университета, задължителен елемент от тип String
- 1.1.1.1.5. id – id на университета, задължителен елемент от тип Integer
- 1.1.1.1.6. regionId – id на региона, задължителен елемент от тип Integer
- 1.1.1.1.7. telephone – телефонен номер на университета, задължителен елемент от тип String
- 1.1.1.1.8. telephoneExtra – допълнителен телефонен номер на университета, незадължителен елемент от тип String
- 1.1.1.1.9. year – година на основаване на университета, задължителен елемент от тип Integer
- 1.1.1.1.10. students – брой студенти в университета (ако не е филиал), незадължителен елемент от тип Integer
- 1.1.1.1.11. site – сайт на университета(ако има такъв), незадължителен елемент от комплексен тип, който е String и представлява референция към сайта на съответния университет
- 1.1.1.2. facultyList –
 - 1.1.1.2.1. faculty – име на факултет, съдържа се неограничен брой пъти в facultyList
- 1.2. branchList – списък на университетите, които имат филиали; може да се съдържа най – много веднъж. Съдържа следните елементи:
 - 1.2.1. branch – университет, който има филиали. Съдържа се неограничен брой пъти в branchList. Съдържа следните елементи:
 - 1.2.1.1. seatName – име на седалището на университета, елемент от тип String
 - 1.2.1.2. uni – празен елемент, който се съдържа точно веднъж в елемента branch. Съдържа атрибути nameRef от тип String и idRef от тип Integer
- 1.3. regionList – списък на регионите в България; може да се съдържа най – много веднъж. Съдържа следните елементи:
 - 1.3.1. regionInfo – име на регион. Съдържа се неограничен брой пъти в regionList. Съдържа следните елементи:
 - 1.3.1.1. regionName – име на региона на университета, елемент от тип String
 - 1.3.1.2. reg – празен елемент, който се съдържа точно веднъж в елемента regionInfo. Съдържа атрибути regRef от тип String и regIdRef от тип Integer

2.3 Тип и представяне на съдържанието

Съдържанието на проекта се намира в един текстов и 19 графични файла:

1 текстов файл – universityCat.xml, който съдържа информацията за университетите, валидиран от XML Schema. Файлът използва UTF-8 кодиране и xml version 1.0.Размер на текстовия файл: 21 KB

19 графични файла – изображения на университетите, получени от техните уеб сайтове. Форматът на изображенията е JPG и всички са преформатирани, за да бъдат с един и същи размер, когато се визуализират в браузър. Общ размер на графичните файлове: 924 KB

3 Дизайн

За цялостното решение на проекта са разработени следните файлове:

- universityCat.xml – основен файл на проекта. Съдържа информация за университетите като използва XML Entities, както и ключове, чрез които се осъществяват връзките между университети и региони, и университети и техните филиали. Документът е валидирам чрез XML Schema
- universityCatScheme.xsd – XML Schema, която съдържа описание на елементите, както и връзките между тях. Примери с описанието на елемента uniInformation, regionList, както и референциите по региони:

```
<xs:element name="uniInformation">
  <xs:complexType>
    <xs:sequence>
      <xs:element name="name" type="xs:string"/>
      <xs:element name="branchName" type="xs:string" minOccurs="0"/>
      <xs:element name="location" type="xs:string"/>
      <xs:element name="region" type="xs:string"/>
      <xs:element name="id" type="xs:integer" />
      <xs:element name="regionId" type="xs:integer"/>
      <xs:element name="telephone" type="xs:string"/>
      <xs:element name="telephoneExtra" type="xs:string" minOccurs="0"/>
      <xs:element name="year" type="xs:integer"/>
      <xs:element name="students" type="xs:integer" minOccurs="0"/>
      <xs:element name="site" minOccurs="0">
        <xs:complexType>
          <xs:simpleContent>
            <xs:extension base="xs:string">
              <xs:attribute ref="href" use="optional"/>
            </xs:extension>
          </xs:simpleContent>
        </xs:complexType>
      </xs:element>
    </xs:sequence>
  </xs:complexType>
</xs:element>
```

Фигура 2: Валидация на елемента uniInformation чрез XML Schema

```

<xs:element name="regionList" maxOccurs="1">
  <xs:complexType>
    <xs:sequence>
      <xs:element name="regionInfo" maxOccurs="unbounded">
        <xs:complexType>
          <xs:sequence>
            <xs:element name="regionName" type="xs:string"/>
            <xs:element name="reg" maxOccurs="1">
              <xs:complexType>
                <xs:simpleContent>
                  <xs:extension base="xs:string">
                    <xs:attribute type="xs:string" name="regRef" />
                    <xs:attribute type="xs:integer" name="regIdRef" />
                  </xs:extension>
                </xs:simpleContent>
              </xs:complexType>
            </xs:element>
          </xs:sequence>
        </xs:complexType>
      </xs:element>
    </xs:sequence>
  </xs:complexType>
</xs:element>

```

Фигура 3: Валидация на елемента regionList чрез XML Schema

```

<xs:key name="regionKey">
  <xs:selector xpath="unisInBG/universityList/university/uniInformation"/>
  <xs:field xpath="region"/>
  <xs:field xpath="regionId"/>
</xs:key>
<xs:keyref name="regionRef" refer="regionKey">
  <xs:selector xpath="unisInBG/regionList/regionInfo/reg"/>
  <xs:field xpath="@regRef"/>
  <xs:field xpath="@regIdRef"/>
</xs:keyref>

```

Фигура 4: Валидация ключа regionKey и референцията regionRef чрез XML Schema

- universityCatXslt.xsl – XSLT файл, чрез който се осъществява трансформацията към HTML, за представяне на съдържанието в уеб браузър
- style.css – css файл, който стилизира съдържанието на проекта
- /images – директория, която съдържа графичните файлове, използвани в проекта



4 Тестване

За тестване на решението на проекта са използвани следните браузъри:

- Mozilla Firefox
- Microsoft Edge



Получена е следната визуализация в гореспоменатите браузъри:

Университети в България

Университети	Филиали	Региони																		
	Софийски университет "Св. Климент Охридски" <table><tr><td>Седалище</td><td>Година</td><td>Тел</td></tr><tr><td>София</td><td>1888</td><td>02/930-8577</td></tr><tr><td>Регион</td><td>Студенти</td><td>Доп.тел</td></tr><tr><td>Югозападен регион</td><td>23769</td><td>02/816-1500</td></tr><tr><td>Сайт</td><td>Факултети</td><td></td></tr><tr><td>www.uni-sofia.bg</td><td></td><td></td></tr></table>		Седалище	Година	Тел	София	1888	02/930-8577	Регион	Студенти	Доп.тел	Югозападен регион	23769	02/816-1500	Сайт	Факултети		www.uni-sofia.bg		
Седалище	Година	Тел																		
София	1888	02/930-8577																		
Регион	Студенти	Доп.тел																		
Югозападен регион	23769	02/816-1500																		
Сайт	Факултети																			
www.uni-sofia.bg																				
	Пловдивски университет "Паисий Хилендарски" <table><tr><td>Седалище</td><td>Година</td><td>Тел</td></tr><tr><td>Пловдив</td><td>1961</td><td>032/261-261</td></tr><tr><td>Регион</td><td>Студенти</td><td>Доп.тел</td></tr><tr><td>Южен централен регион</td><td>17825</td><td>032/261-303</td></tr></table>		Седалище	Година	Тел	Пловдив	1961	032/261-261	Регион	Студенти	Доп.тел	Южен централен регион	17825	032/261-303						
Седалище	Година	Тел																		
Пловдив	1961	032/261-261																		
Регион	Студенти	Доп.тел																		
Южен централен регион	17825	032/261-303																		



Фигура 5: Изглед на проекта, тестван с Microsoft Edge. Университети

Университети в България

Университети	Филиали	Региони
Технически университет - София		
	Филиал Пловдив	
	Седалище	Тел
	Пловдив	032/659-559
	Регион	Факултети
	Южен централен регион	Сайт
	Филиал Сливен	
	Седалище	Тел
	Сливен	044/667-552
	Регион	Доп.тел
	Югоизточен регион	Сайт
	Филиал Сливен	
	Седалище	Тел
	Сливен	044/667-552
	Регион	Доп.тел
	Югоизточен регион	Сайт

Фигура 6: Изглед на проекта, тестван с Microsoft Edge. Филиали

Университети в България

Университети	Филиали	Региони	
Северозападен регион			
	Медицински университет - Плевен		
	Седалище	Тел	Факултети
	Плевен	064/884-101	
	Година	Доп.тел	
	1974	064/800-728	
Студенти	Сайт		
1458	www.mu-pleven.bg		
Северен централен регион			
	Русенски университет "Ангел Кънчев"		
	Седалище	Тел	Факултети
	Русе	082/888-465	
	Година	Доп.тел	

Фигура 7: Изглед на проекта, тестван с Microsoft Edge. Региони

5 Заключение и възможно бъдещо развитие

С използването на езика XML, проектът е добре структуриран и връзките между елементите са представени ясно и разбираемо. Чрез XSLT се осъществява лесно преобразуването, за да бъде достъпно съдържанието през уеб браузър. При наличие на промяна, в зависимост от това каква е тя, промени трябва да бъдат извършени на по-малко места, тъй като съдържанието и оформлението са разделени в отделни файлове.

В посока на бъдещото развитие на проекта, могат да бъдат включени всички университети в България, както и специалностите, които техните филиали предлагат. Могат да бъдат включени и рейтинги на университетите.

6 Разпределение на работата

Както документите по проекта – .xml, .xslt, .xsd, .css, така и документацията, са разработени съвместно от членовете на екипа.

7 Използвани литературни източници и Уеб сайтове

1. www.uni-sofia.bg - Уеб сайт на Софийски университет "Св. Климент Охридски"
2. www.tu-sofia.bg - Уеб сайт на Технически университет - София
3. www.tu-plovdiv.bg - Уеб сайт на Технически университет – София - Филиал Пловдив
4. www.tu-sliven.com - сайт на Технически университет – София - Филиал Сливен
5. www.natfiz.bg - Уеб сайт на Национална академия за театрално и филмово изкуство "Кръстьо Сарафов"
6. www.aubg.edu - Уеб сайт на Американски университет в България
7. www.uni-plovdiv.bg - Уеб сайт на Пловдивски университет "Паисий Хилендарски"
8. www.uni-kardzhali.com - Уеб сайт на Пловдивски университет "Паисий Хилендарски" - Филиал Кърджали
9. www.uni-sz.bg - Уеб сайт на Тракийски университет Стара Загора
10. www.shu-bg.net - Уеб сайт на Шуменски университет "Епископ Константин Преславски"
11. www.k-dobrich.shu.bg - Уеб сайт на Шуменски университет "Епископ Константин Преславски" - Филиал Добрич
12. www.naval-acad.bg - Уеб сайт на Висше военноморско училище "Никола Вапцаров"
13. www.uni-ruse.bg - Уеб сайт на Русенски университет "Ангел Кънчев"
14. www.mu-pleven.bg - Уеб сайт на Медицински университет – Плевен
15. Упражнения по "XML технологии за семантичен Уеб", зимен семестър 2017/2018, проф. Д-р Боян Бончев
16. Лекции по "XML технологии за семантичен Уеб", зимен семестър 2017/2018, проф. Д-р Боян Бончев